Published in final edited form as:

Int J Parasitol. 2008 August ; 38(10): 1111–1119. doi:10.1016/j.ijpara.2008.02.002.

Characterization of the promoter of the *Brugia malayi* 12 kDa small subunit ribosomal protein (RPS12) gene

Ana de Oliveira^a, Charles R. Katholi^b, and Thomas R. Unnasch^{c,*}

^a Gorgas Center for Geographic Medicine, University of Alabama at Birmingham, USA

^b Department of Biostatistics, University of Alabama at Birmingham, USA

^c Department of Global Health, University of South Florida, USA

Abstract

Unravelling gene regulatory mechanisms in human filarial parasites will require an understanding of their basic promoter structure. Only a single promoter from a human filarial parasite has been characterised in detail, the 70 kDa heat shock promoter of *Brugia malayi* (BmHSP70). This promoter was found to lack features found in a typical eukaryotic promoter. To determine if this was unique to the BmHSP70 promoter, a detailed analysis was undertaken of the promoter for the *B. malayi* small subunit 12 kDa ribosomal protein (BmRPS12) gene. The BmRPS12 promoter contained a unique tandem repeat structure. Deletion of these repeats resulted in the loss of 80% of promoter activity. Block replacement mutagenesis identified five regions outside the repeat which were essential for promoter activity. No predicted binding sites for proteins that normally associate with the typical eukaryotic core promoter domains were found in the essential domains or the repeat region. However, the repeat region contained many putative binding sites for GATAA transcription factor family proteins. Of 20 upstream domains of other ribosomal protein genes, one contained a repeat structure similar to that found in the BmRPS12 promoter, and the majority encoded putative GATAA transcription factor binding sites. This study demonstrates that the BmRPS12 promoter, like the BmHSP70 promoter, is distinct from a typical eukaryotic promoter.

Keywords

Filariasis; Promoter; Gene regulation; Transfection; Biologistics

1. Introduction

Filarial parasites represent a significant public health problem worldwide. It is estimated that over 140 million individuals are infected with filarial parasites, with 40 million of these suffering disfigurement or incapacitation (World Health Organization, 1995, 2000). Research efforts spearheaded by the Filarial Genome Project and The Institute for Genome Research (TIGR) have generated a significant body of knowledge concerning the genes expressed by human filarial parasites, and in particular *Brugia malayi* (Blaxter et al., 2002). These efforts have culminated in a sequence database representing nine-fold coverage of the *B. malayi* genome (Ghedini et al., 2007).

*Corresponding author. Address: University of South Florida, College of Public Health, Department of Global Health, Global Health Infectious Disease Research Program, 3720 Spectrum Blvd., Suite 304, Tampa, FL 33612, USA. Tel.: +1 813 974 0507; fax: +1 813 974 0992. E-mail address: tunnasch@health.usf.edu (T.R. Unnasch).

The advent of the *B. malayi* genome sequence will open up new avenues of research into how gene expression in filarial parasites is regulated. This question will be central to understanding how this parasite has adapted to life in two very different host environments (vertebrate and insect) and how it survives in the face of an ongoing attack by the host's immune system. With the appropriate tools, this information also has the potential to directly link gene polymorphisms to particular phenotypes (e.g. drug resistance).

To date, little is known about how filarial parasites regulate gene expression, partly due to difficulties in carrying out genetic studies in these organisms. For example, the obligate parasitic life cycle of these organisms has made it impossible to perform conventional genetic studies, as one cannot isolate mutants with easily scored phenotypes, and it is impractical to carry out defined genetic crosses. In the absence of classical genetics, reverse genetic approaches can be used to study gene function and regulation. In recent years, substantial progress has been made in adapting reverse genetic approaches to the study of human filarial parasites. For example, RNA interference (RNAi) (Fire et al., 1998) has been shown to be capable of reducing gene expression in both *B. malayi* (Aboobaker and Blaxter, 2003) and in the related human filarial parasite *Onchocerca volvulus* (Lustigman et al., 2004). Furthermore, studies have demonstrated *B. malayi* can be transiently transfected using both biolistic and microinjection methods (Higazi et al., 2002). Biolistic transfection has been used to explore both promoter structure and trans-splicing in *B. malayi* (Higazi et al., 2002, 2005; Shu et al., 2003; Higazi and Unnasch, 2004; Liu et al., 2007). These studies have suggested that the promoter structure of this parasite is relatively unique. For example, detailed mapping studies of the promoter of the *B. malayi* 70 kDa heat shock protein (BmHSP70) revealed that this promoter contained four essential domains, ranging in size from 6 to 22 nucleotides (nt) (Higazi et al., 2005). The two most distal domains encoded a binding site for the heat shock transcription factor and a putative binding site for a GAGA transcription factor, motifs that are found in many other HSP70 promoters. However, none of these essential domains contained sequences found in the core domain of a typical eukaryotic promoter, such as CAAT or TATAA boxes (Higazi et al., 2005). The largest essential domain was located at positions -53 to -32 relative to the start of the open reading frame (ORF) and included the splice leader (SL) addition site. The activity of this domain was not related to SL addition, as transgenic transcripts produced from *B. malayi* transfected with constructs containing the BmHSP70 promoter alone were not trans-spliced (Shu et al., 2003). These data suggest that the regulatory domains of the BmHSP70 promoter were similar to those found in other eukaryotes, but that the core promoter domain exhibited features that appeared to be distinct from those found in most other well-characterised eukaryotic promoters. An analysis of two additional promoters of two *B. malayi* highly transcribed genes, the first encoding the *B. malayi* homologue of the 12 kDa peptide of the small subunit of the ribosome (the BmRPS12 gene) and the second encoding a putative RNA binding protein (the BmRBP1 gene) demonstrated that both were active in promoting transcription in the transient transfection system, whilst also lacking features commonly found in most eukaryotic core promoters (Higazi et al., 2005). Together, these studies suggested that the core domains of *B. malayi* promoters may lack many of the conserved elements found in most eukaryotic promoters. However, it has not been possible to identify what the conserved domains of the *B. malayi* promoters actually are, as to date only the BmHSP70 promoter has been mapped in detail. In the results presented below, we report the detailed mapping of a second *B. malayi* promoter, derived from the BmRPS12 gene. Together with the previous studies reporting the detailed mapping of the BmHSP70 promoter, these data provide further insight into the structural features that may represent the core promoter domains of this parasite. These data also provide some insight into the transcriptional regulators that may be important in controlling expression of the BmRPS12 gene, and possibly other ribosomal protein genes of this organism.

2. Materials and methods

2.1. Preparation of deletion and substitution mutants of the BmRPS12 promoter

The parental clone of the promoter for the BmRPS12 gene (BmRPS12(−641 to −1)/luc) was prepared in the reporter vector pGL3 basic, as previously described (Higazi et al., 2005). Deletions of the 5' domain of the parental construct and the repeat domains were prepared by restriction enzyme mediated inverse PCR, essentially as previously described (Shu et al., 2003). This process resulted in the replacement of the region to be deleted with a SpeI restriction site. The 12 nt substitution mutants were prepared using a similar procedure (Higazi et al., 2005). In this case, outward facing primers were spaced 12 nt apart in the parental sequence. The primers were designed to contain a SpeI site at their 5' ends followed by 3 nts that differed from those in the parental sequence. These primers were used in an inverse PCR employing the parental construct as a template. The resulting amplicon was digested with SpeI, gel purified, self-ligated and used to transform *Escherichia coli*. This resulted in the replacement of the 12 nt region in question by an SpeI site flanked on each side by the three mutated nts. The 3–4 nt substitution mutants were prepared using appropriate mutated oligonucleotides and using the GeneTailor site-directed mutagenesis system (Invitrogen, Carlsbad, CA), following the manufacturer's protocol. In developing the sequences for the irrelevant triplets flanking the SpeI site in the 12 nt replacements and the mutant sequences in the 3–4 nt replacements, each base to be mutated was replaced by the one that was most distantly related to it. Thus, A was replaced with C, G with T, C with A and T with G. The DNA sequences of all constructs were determined to ensure that the mutated sequences were as expected and that the amplification process had not introduced any unwanted additional mutations.

2.2. Transient transfection and analysis of promoter activity

Isolated *B. malayi* embryos were transfected and promoter activity assayed by luciferase activity as previously described (Shu et al., 2003). In brief, embryos were isolated from gravid female parasites and transfected with the experimental DNA driving the expression of firefly luciferase mixed with a constant amount of an internal standard, consisting of the BmHSP70 promoter fragment driving the expression of renilla luciferase (construct BmHSP70(−659 to −1)/ren) (Shu et al., 2003). Transfected embryos were maintained in culture for 48 h before being assayed for transgene activity. Firefly luciferase activity was normalised to the amount of renilla luciferase activity in each sample to control for variations in transfection efficiency. Firefly/renilla activity ratios for each sample were further normalised to the activity ratio found in embryos transfected in parallel in each experiment with the parental construct (BmRPS12(−641 to −1)/luc). This permitted comparisons of data collected in experiments carried out on different days. Each construct was tested in at least two independent experiments, with each experiment containing triplicate transfections of each construct to be analysed. The statistical significance of differences in the activity of the experimental constructs was determined using Dunnett's test, as previously described (Shu et al., 2003).

2.3. Bioinformatic analyses

2.3.1. Analysis of putative transcription factor binding sites

The four regions of the BmRPS12 promoter domain found to be important in promoter activity (−437 to −380, −343 to −340, −63 to −37 and the repeat domain) were analysed separately for the presence of sequences similar to those known to bind transcription factor families using the transcription element search system (TESS) [Schug, J., Overton, G.C., 1997. TESS: Transcription Element Search Software on the WWW., Computational Biology and Informatics Laboratory, School of Medicine, University of Pennsylvania, Philadelphia, PA]. In the case of the −343 to −340 domain, the sequence analysed included the nts extending from −353 to −334, in order to include at least 20 nt in the analysis. For the repeat domain, a single

repeat unit (containing the consensus sequence for the repeat) was analysed. The programme was run with its default values, with the exception that the cutoff P value was set at 1×10^{-3} .

2.3.2. Identification and analysis of upstream domains of genes encoding ribosomal proteins

A text search was first performed of the NemBase2 expressed sequence tag (EST) database (www.nematodes.org/nematodeESTs/nembase.html) with the terms “ribosome OR ribosomal” to identify ESTs encoding other ribosomal proteins. A total of 44 ESTs encoding putative ribosomal protein homologues were identified in this manner, of the approximately 50 proteins expected to be present in the *B. malayi* ribosome (Alberts et al., 2002). These 44 ESTs were then further analysed to identify those that contained a putative full-length mRNA sequence. EST sequences were judged to be full-length if they contained a putative full-length ORF (as determined by homology of the derived amino acid sequence to the amino terminus of the derived amino acid sequence of its predicted homologue), together with an SL located upstream from the start of the full-length ORF. Of the 44 ESTs identified, 21 met these criteria. The 1 kbp upstream of the SL addition sites of the genes were then abstracted from the draught *B. malayi* genome sequence produced by the Institute for Genome Research (TIGR) (Ghedini et al., 2007) using BLASTn searches employing the 5' 100 nt of each EST as the query sequence.

The upstream domains were examined for the presence of repeat elements by conducting a matrix similarity search comparing each sequence to itself. In this analysis, the entire 1 kbp upstream of the start of the ORF were analysed. The comparison was conducted with a query sequence length of 10 nt and a minimal match of 80%. Any sequences exhibiting off diagonal matches were then examined manually in the region of interest to identify repeated sequences. The 500 nt extending upstream of the start codon of each upstream domain were analysed for the presence of GATAA factor binding sites using TESS, as described above.

3. Results

3.1. Preliminary mapping of the BmRPS12 promoter

Previous studies had demonstrated that the 886 nt upstream of the start of the BmRPS12 ORF were capable of driving transcription of the luciferase reporter gene in the *B. malayi* transient transfection system (Higazi et al., 2005). This region contained two notable features (Fig. 1A). The first was the presence of a core promoter-like domain, which was predicted by analysis of a core promoter prediction algorithm trained on *Drosophila melanogaster* promoter sequences (www.fruitfly.org/cgi-bin/seq_tools/promoter.pl). This putative core promoter domain extended from positions -748 to -696. However, a 5' deletion removing this putative core domain (BmRPS12 -688 to -1/luc) had no effect on the level of luciferase reporter gene activity, suggesting that it was not playing a role in the promoter function of this sequence (Higazi et al., 2005). The second notable feature of the upstream domain was the presence of a repeat element. This element extended from positions -318 to -65 and consisted of 5 3/4 almost exact repeats of a 44 nt sequence (Fig. 1A). This repeat domain ended just 10 nt upstream of the SL addition site found in the native gene (Fig. 1A).

As a first step in mapping the essential domains of the BmRPS12 promoter, a series of 70 nt 5' deletions were constructed, beginning with construct BmRPS12 (-688 to -1)/luc, which as discussed above, previous studies had demonstrated produced levels of promoter activity that were identical to those obtained from BmRPS12 (-886 to -1)/luc, the construct shown schematically in Fig. 1A. These constructs were tested for promoter activity as described in Section 2. Deletion of nts from -688 to -463 resulted in no changes in promoter activity (Fig. 1B). However, deletion of any nts downstream from this point resulted in a dramatic decrease

in promoter activity. This localised the core promoter of the BmRPS12 gene to a region extending 462 nt upstream of the start of the ORF. This region contained the SL addition site, the repeat domain and the 144 nt extending upstream from the repeat domain.

3.2. Role of the repeat domain in the BmRPS12 promoter's activity

Because the repeat domain represented over half of the core promoter domain identified by the 5' deletion analysis, it was of interest to determine if this region played a role in the promoter's activity, and if so, if all of the repeat domains were necessary to obtain full activity. To answer this question a series of internal deletions of BmRPS12 (−688 to −1)/luc were prepared, sequentially removing increasing numbers of the repeat family units. These deletions were then tested for promoter activity. Constructs containing 4 3/4, 3 3/4 or 2 3/4 repeat units exhibited promoter activities that were statistically indistinguishable from the parental construct (containing 5 3/4 repeats; Fig. 2). Constructs containing 1 3/4 repeats or 3/4 of a repeat unit exhibited levels of promoter activity that were reduced approximately 40% from those seen with the parental construct (Fig. 2). Deletion of all repeat units resulted in a construct that exhibited a promoter activity that was reduced roughly 80% from that seen with the parental construct (Fig. 2).

The finding that a repeated sequence element was part of the BmRPS12 promoter raised the question whether similar structures were found in other *B. malayi* promoters. As a first step in answering this question, the BmRPS12 repeat sequence was used as a query sequence to conduct a BLAST search of the TIGR *B. malayi* whole genome sequence (WGS) database and the Genbank complete nucleotide sequence collection. Apart from the BmRPS12 gene itself, no significant homologies were detected in these searches. A search was then made to determine if other *B. malayi* putative promoter elements might contain a similar repeat element structure, but utilising repeat elements distinct from those in the BmRPS12 gene. To accomplish this, the upstream domains of 20 additional genes encoding ribosomal proteins were analysed for the presence of repeat units. Of the 20 sequences analysed, one (the upstream domain of the putative BmRPS20 gene) contained a similar repeated sequence element. This element consisted of a repeat structure containing 18 repeat units with a unit repeat length of 30 nt. This repeat element was located at positions −1034 to −500 upstream of the start of the ORF, and −946 to −412 relative to the SL addition site.

3.3. Identification of essential domains in the BmRPS12 promoter

These results suggested that the repeat domains were important in regulating the level of transcription from the BmRPS12 promoter. However, because deletion of all repeats resulted in a construct that still exhibited an activity level that was still roughly 20% of that of the parental construct, elements essential for maximal promoter function must reside outside the repeat domain. To identify these elements, a series of 12–14 nt block replacements spanning the segments located 5' and 3' of the repeat domain within the −462 to −1 region were prepared. These constructs were then tested for activity in the transient transfection system. These studies localised the essential domains of the BmRPS12 promoter to three areas flanking the repeat domain, at positions −438 to −381, −344 to −333 and −64 to −38 (Fig. 3A). Replacement of these three regions resulted in decreases of promoter activity ranging from 75% to 99% (Fig. 3A). These regions were mapped in more detail using a panel of 3–4 nt block replacements (Fig. 3B). These studies revealed that the −438 to −381 region contained two important sub-domains at positions −438 to −419 and −410 to −381. The −344 to −333 region contained a single important sub-domain at positions −344 to −341. The −64 to −38 domain contained two important subdomains at positions −64 to −58 and −49 to −42. Promoter activity was completely lost when portions of the −438 to −419, −410 to −381 and −344 to −341 sub-domains were replaced (Fig. 3B), whilst a greater than 90% reduction in promoter activity was observed upon replacement of the −61 to −58 region.

The important sub-domains identified by the block replacement mutagenesis experiments are summarised in Fig. 4. Three sub-domains were located 5' to the repeat region, whilst two were located 3' to the repeats. Interestingly, the two sub-domains 3' to the repeat domain flanked, but did not include, the SL addition site (Fig. 4). This finding was similar to what was found for the BmHSP70 promoter, where the largest essential domain was found to include the SL addition site (Higazi et al., 2005). A comparison of the sequences flanking the SL addition sites of the BmHSP70 and BmRPS12 promoters indicated that these domains shared several common residues located in the regions found to be essential for promoter activity. These included a polypyrimidine tract located 5' to the SL addition site and a triplet conserved sequence located 3' to the SL addition site (Fig. 5).

3.4. Bioinformatic analysis of the BmRPS12 promoter

The essential domains of the BmRPS12 promoter and the repeat unit sequences were then analysed to determine whether they contained sequences that might bind known transcription factors, using the TESS search engine as described in Section 2. The region encompassing positions -354 to -335 did not contain any sequences known to interact with transcription factors (Table 1). This was also true of the -64 to -42 region, which contained the SL addition site (Table 1). This was similar to what was found for the region surrounding the SL addition site of the BmHSP70 promoter, which also does not encode sequences that are predicted to interact with any known transcription factor (Higazi et al., 2005). The -438 to -381 region was found to encode a sequence similar to the binding site for members of the CTF/NF1 family of transcription factors (Table 1).

In contrast to the essential domains located outside the repeat domain, the BmRPS12 repeat unit was found to contain several potential binding sites for putative GATAA transcription factors (Table 1). Each repeat unit contained four such putative binding GATAA factor binding sites (Fig. 6A), leading to a total of 22 such putative elements in the entire BmRPS12 repeat array. The large number of GATAA homologues in the BmRPS12 repeat unit suggested that a factor binding this sequence might be involved in the regulation of transcription of the BmRPS12 gene. As the expression of the ribosomal proteins is expected to be coordinately regulated (Li et al., 1999), it was of interest to determine whether similar sites might exist in other ribosomal protein genes. The 500 nt extending upstream of each of the 20 ribosomal protein genes described above were therefore searched for similar motifs, as described in Section 2. Of the 21 putative promoters for the ribosomal proteins analysed (the BmRPS12 promoter plus the additional 20 putative promoters described above), all but one were found to contain putative GATAA factor binding sites, with a large proportion containing multiple examples of this motif (Fig. 6B).

4. Discussion

Very little is known concerning the promoter structure of the genes of the human filarial parasites. To date, the essential domains of a single promoter from these parasites (the BmHSP70 promoter) have been mapped in detail (Higazi et al., 2005). The BmHSP70 promoter was found to lack many features, such as CAAT and TATAA boxes, which are normally found in the core domains of eukaryotic promoters. The data presented above demonstrate that these deficiencies are shared by the BmRPS12 promoter. Similar to what was noted in the case of the BmHSP70 promoter, the essential domains of the BmRPS12 promoter lacked identifiable binding sites for many of the transcription factors that together make up the pre-initiation complex which forms at the core domain of typical eukaryotic promoters. Thus, both promoters seem to lack essential domains that bind to the TATA box binding protein (TBP) and transcription factor IID (TFIID). Despite this, predicted homologues of these conserved transcription factors appear to be present amongst the predicted genes in the *B. malayi* genome

(data not shown). An analysis of the DNA sequence binding specificity of these homologues would be useful in determining if these proteins interact with typical CAAT and TATA box-like motifs, or if they recognise less typical sequences.

The BmRPS12 promoter was found to contain three domains which, when mutated, resulted in the loss of over 90% of promoter activity. The largest of these was located in positions -438 to -381 relative to the start codon, and could be divided into two sub-domains. A second small domain was located at positions -344 to -341. The final domain was located at positions -64 to -42. It also could be divided into two sub-domains. This structure was similar in some respects to the arrangement of essential domains found in the BmHSP70 promoter. Similar to the BmRPS12 promoter, three domains in the BmHSP70 promoter were identified which, when mutated, resulted in the loss of over 90% of promoter activity. A comparison of the sequences most of the essential domains of the BmHSP70 and BmRPS12 promoters revealed no striking similarities (data not shown). However, mutation of the regions flanking the SL addition site in both promoters resulted in greater than 90% reductions in promoter activity. These domains also exhibited some degree of sequence conservation, and at least some of the conserved residues were located in regions that were essential for activity in both promoters. The role of these sequences in regulating transcription appears to be distinct from their function as SL acceptors, for two reasons. First, neither promoter's actual SL acceptor site was essential for activity, as mutation of these residues did not result in elimination of luciferase expression. Second, previous studies have demonstrated that sequences downstream of the start of the ORF are necessary for trans-splicing from transgenic transcripts produced from both of these genes (Liu et al., 2007) and constructs lacking these downstream motifs (such as those analysed here) produce transgenic mRNAs that are not trans-spliced (Liu et al., 2007). These data suggest a role for control of expression of the sequences flanking the SL addition site in transcription which is distinct from the role played in trans-splicing, suggesting that the region surrounding the SL addition site may encode a portion of the core promoter in both of these genes. It is tempting to speculate that the regions that were conserved in both promoters and which were also essential for promoter activity may encode the putative core promoter domains. However, it is possible that some of these conserved sequences may also be involved in the trans-splicing process. Further work will be required to separate these processes.

The most striking feature of the BmRPS12 promoter is the presence of the 44 nt repeat unit. Promoter constructs with 2 3/4 repeat units exhibited the same levels of promoter activity as those with the full complement of 5 3/4 repeats, indicating that all repeat units were not necessary for full promoter activity. In contrast, constructs with 1 3/4 or 3/4 of a repeat unit exhibited activity levels that were roughly half that of constructs with at least 2 3/4 repeats, and deletion of all of the repeats resulted in the loss of about 80% of the promoter activity. These data suggest the hypothesis that the repeat units encoded a recognition site for a factor that was important in regulating BmRPS12 promoter activity. Furthermore, the fact that each repeat unit contained multiple copies of a GATAA transcription factor recognition site-like motif suggests that several molecules of the putative regulatory factor might bind to the repeat unit. Tandemly repeated copies of such transcriptional regulatory factors have been seen in other eukaryotic promoters, and often result in cooperative binding of multiple copies of the transcriptional regulator to the promoter in question. For example, most eukaryotic heat shock promoters encode multiple heat shock elements to which heat shock transcription factors bind cooperatively, resulting in a several fold increase in transcription from the associated promoter (Amin et al., 1994).

The repeat units present in the BmRPS12 promoter appear to be relatively unique, since no other homologues of the repeat unit could be found in the *B. malayi* genome. Furthermore, similar repeat structures were found in only one of the additional 20 upstream domains examined from other ribosomal protein genes. However, homologues of the GATAA motif

were found in the majority of the ribosomal protein upstream domains examined. This suggests that the putative factors that interact with the BmRPS12 repeat unit may also interact with the upstream domains of many of the other ribosomal protein genes. Transcription factors recognising the GATAA motif are members of the zinc finger family (He et al., 2007). Putative zinc finger proteins are very numerous in the free living nematode *Caenorhabditis elegans* representing approximately 3% of the genome (Clarke and Berg, 1998). The genome of *B. malayi* contains 276 genes that are predicted to encode zinc finger proteins, of which four are predicted to be members of the GATA transcription factor family (data not shown). It is tempting to hypothesise that one of these putative transcription factors interacts with the GATAA motifs in the BmRPS12 repeat unit and is responsible for regulating transcription from this gene, as well as the other members of the ribosomal protein family whose promoters contain similar motifs. Experiments testing this hypothesis are currently underway.

Acknowledgements

We would like to thank Dr. Naomi Lang-Unnasch for critical reading of the manuscript prior to publication and Dr. Julian Rayner for critical reading and helpful discussions. We also acknowledge the use of the *B. malayi* genomic sequence produced and maintained by The Institute for Genome Research (TIGR). TIGR's sequencing effort was part of the International Brugia Genome Sequencing Project and was supported by an award from the National Institute of Allergy and Infectious Diseases, National Institutes of Health. This work was supported by a grant from National Institute of Allergy and Infectious Diseases to T.R.U. (Project No. R01AI048562).

References

- Aboobaker AA, Blaxter ML. Use of RNA interference to investigate gene function in the human filarial nematode parasite *Brugia malayi*. *Molecular and Biochemical Parasitology* 2003;129:41–51. [PubMed: 12798505]
- Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*. Garland Science; New York: 2002.
- Amin J, Fernandez M, Ananthan J, Lis JT, Voellmy R. Cooperative binding of heat shock transcription factor to the Hsp70 promoter *in vivo* and *in vitro*. *Journal of Biological Chemistry* 1994;269:4804–4811. [PubMed: 8106450]
- Blaxter M, Daub J, Guiliano D, Parkinson J, Whitton C. The *Brugia malayi* genome project: expressed sequence tags and gene discovery. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2002;96:7–17. [PubMed: 11925998]
- Clarke ND, Berg JM. Zinc fingers in *Caenorhabditis elegans*: finding families and probing pathways. *Science* 1998;282:2018–2022. [PubMed: 9851917]
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998;391:806–811. [PubMed: 9486653]
- Ghedini E, Wang S, Spiro D, Caler E, Zhao Q, Crabtree J, Allen JE, Delcher AL, Guiliano DB, Miranda-Saavedra D, Angiuoli SV, Creasy T, Amedeo P, Haas B, El-Sayed NM, Wortman JR, Feldblyum T, Tallon L, Schatz M, Shumway M, Koo H, Salzberg SL, Schobel S, Pertea M, Pop M, White O, Barton GJ, Carlow CK, Crawford MJ, Daub J, Dimmic MW, Estes CF, Foster JM, Ganatra M, Gregory WF, Johnson NM, Jin J, Komuniecki R, Korf I, Kumar S, Laney S, Li BW, Li W, Lindblom TH, Lustigman S, Ma D, Maina CV, Martin DM, McCarter JP, McReynolds L, Mitreva M, Nutman TB, Parkinson J, Peregrin-Alvarez JM, Poole C, Ren Q, Saunders L, Sluder AE, Smith K, Stanke M, Unnasch TR, Ware J, Wei AD, Weil G, Williams DJ, Zhang Y, Williams SA, Fraser-Liggett C, Slatko B, Blaxter ML, Scott AL. Draft genome of the filarial nematode parasite *Brugia malayi*. *Nature* 2007;317:1756–1760.
- He C, Cheng H, Zhou R. GATA family of transcription factors of vertebrates: Phylogenetics and chromosomal synteny. *Journal of Biosciences* 2007;32:1273–1280. [PubMed: 18202451]
- Higazi TB, Merriweather A, Shu L, Davis R, Unnasch TR. *Brugia malayi*: transient transfection by microinjection and particle bombardment. *Experimental Parasitology* 2002;100:95–102. [PubMed: 12054699]
- Higazi TB, Unnasch TR. Intron encoded sequences necessary for trans splicing in transiently transfected *Brugia malayi*. *Molecular and Biochemical Parasitology* 2004;137:181–184. [PubMed: 15279965]

- Higazi TB, DeOliveira A, Katholi CR, Shu L, Barchue J, Lisanby M, Unnasch TR. Identification of elements essential for transcription in *Brugia malayi* promoters. *Journal of Molecular Biology* 2005;353:1–13. [PubMed: 16154590]
- Li B, Nierras CR, Warner JR. Transcriptional elements involved in the repression of ribosomal protein synthesis. *Molecular and Cellular Biology* 1999;19:5393–5404. [PubMed: 10409730]
- Liu C, de Oliveira A, Higazi TB, Ghedin E, DePasse J, Unnasch TR. Sequences necessary for trans-splicing in transiently transfected *Brugia malayi*. *Molecular and Biochemical Parasitology* 2007;156:62–73. [PubMed: 17727976]
- Lustigman S, Zhang J, Liu J, Oksov Y, Hashmi S. RNA interference targeting cathepsin L and Z-like cysteine proteases of *Onchocerca volvulus* confirmed their essential function during L3 molting. *Molecular and Biochemical Parasitology* 2004;138:165–170. [PubMed: 15555728]
- Shu L, Katholi C, Higazi T, Unnasch TR. Analysis of the *Brugia malayi* HSP70 promoter using a homologous transient transfection system. *Molecular and Biochemical Parasitology* 2003;128:67–75. [PubMed: 12706798]
- World Health Organization. Fact Sheet #102. World Health Organization; Geneva: 2000. Lymphatic filariasis.
- World Health Organization. WHO Technical Report #852. World Health Organization; Geneva: 1995. Onchocerciasis and its control.

Fig. 1.

Terminal deletion analyses of the BmRPS12 promoter. (A) Schematic diagram of the major features of the region upstream of the BmRPS12 open reading frame (ORF). Coordinates refer to the position relative to the start of the ORF (indicated by ATG on the diagram). Core, the region predicted to encode a putative core promoter domain, as described in the text; Repeats, repeat domain; SL, spliced leader addition site. (B) Activity of nested 5' deletions of the BmRPS12 promoter in transiently transfected *Brugia malayi* embryos. Bars represent the means of at least six independent transfections. Error bars represent the SDs of the independent transfections. Asterisks indicate constructs whose activity was significantly different from the control ($P < 0.05$).

Fig. 2.

Promoter activity of constructs with varying numbers of the 44 nucleotide (nt) repeat unit. Constructs with varying numbers of the 44 nt repeat were prepared and tested for promoter activity as described in the text. The bars represent mean values derived from at least six independent transfections, whilst the error bars indicate the SDs surrounding the means. The letters indicate groups of constructs in which the promoter activities of the members of the group were not significantly different from one another ($P > 0.05$).

Fig. 3.

Analysis of the promoter activity of block replacement mutants of the BmRPS12 promoter. (A) Promoter activity of 12–14 block replacements of the regions located 5' and 3' of the repeat domain. Grey bars indicate means and SDs of at least six independent transfections. The black separator indicates the position of the repeat domain. Asterisks indicate constructs whose activities were significantly lower than that seen in embryos transfected with the wild type promoter ($P < 0.01$). (B) Promoter activity of 3–4 nucleotide block replacements of the domains found to represent essential domains by the experiments presented in (A). The bars represent means and SDs of the promoter activity of each mutant. Each mutant was tested in at least six independent transfections. Similarly shaded bars represent constructs derived from each of the

three essential domains identified in the experiments presented in (A). Asterisks indicate constructs whose activity was significantly lower than that seen in embryos transfected with the wild type promoter.


```

-455      -445      -435      -425      -415
      GCTTAAGG AGAATTTTAA AAAAACTATA GAGCTACAAA ATACGACAGT
-405      -395      -385      -375      -365      -355
TTTTGAAGCT GGCAATATGC CAATAAAGCA TAGCACTTTT GGACTTTTGG GTAGAATTG
-345      -335      -325      -315      -305      -295
GACATGGATG GTTGTTTATG TGTGGGCAAT GGTGATTGA TTAGCTGATA ATTGATATGA
-285      -275      -265      -255      -245      -235
TAATTTTCGAT GCTGTTTAGG TTGATTAGCT GATAATTGAT ATGATAATTT CGATGCTGTT
-225      -215      -205      -195      -185      -175
TAGGTTGATT AGCTGATAAT TGATATGATA ATTTCGATGC TGTTTAGGTT GATTAGCTGA
-165      -155      -145      -135      -125      -115
TAATTGATAT GATAATTTTCG ATGCTGTTTA GGTGATTGG CTGATAATTG ATATGATAAT
-105      -95      -85      -75      -65      -55
TTCAGTGCTG TTTAGGTTGA TTAGCTGATA ATTGATATGA TAATTTCAGT ATTCCTTTAG
-45      -35      -25      -15      -5
GTTTGATTAG TTGGTAATTG ATAGAGGTGC ACTCATCGAG GACAACTGCT GAACATG
↑

```

Fig. 4.

Domains essential for promoter activity in the BmRPS12 promoter. Sequences highlighted in grey indicate regions which, when mutated, resulted in the loss between 50% and 90% of promoter activity. Sequences highlighted in black represent regions which, when mutated, resulted in a loss of >90% of promoter activity. The alternating single and double underlining highlights the individual repeat units in the repeat domain. Within the repeat domain, positions that differ from the repeat consensus sequence are highlighted in bold type. The arrow below the text indicates the splice leader addition site. The initiating codon of the open reading frame is indicated by italic type.

Fig. 5.

Comparison of the regions flanking the spliced leader addition site in the BmHSP70 and BmRPS12 promoters. Sequences highlighted in grey indicate regions which, when mutated, resulted in the loss between 50% and 90% of promoter activity. Sequences highlighted in black represent regions which, when mutated, resulted in a loss of >90% of promoter activity. The arrow indicates the position of the splice leader addition site.

Fig. 6.

Putative GATAA transcription factor binding sites in the BmRPS12 repeat unit and other ribosomal promoters. (A) GATAA transcription factor binding sites in the BmRPS12 repeat. Sequences similar to the GATAA transcription factor binding site are indicated by underlining. Within each putative binding element, the nucleotides matching the factor consensus sequence are indicated by large type. (B) Representation of GATAA elements in other ribosomal protein promoters. The number of GATAA elements in the upstream element examined is given on the x axis and the number of genes containing that number of GATAA elements is given in the y axis. GATAA elements were identified as described in Section 2.

Table 1

Predicted transcription factor binding sites in the regions necessary for promoter activity

Region	Factor	Recognition site	Sequence model ^a	<i>P</i> value ^b
–438 to –381	CTF/NF1	–405 to –392	TGGCNNNNNGCCAA	9.2×10^{-4}
–354 to –335	None			
Repeat	GATA-1	Multiple	WGATAA	4×10^{-4}
–64 to –42	None			

^a Consensus sequence of the identified binding site.

^b Poisson *P* value.